

Constructive sparse trigonometric approximation and other problems for functions with mixed smoothness*

V. Temlyakov[†]

December 31, 2014

Abstract

Our main interest in this paper is to study some approximation problems for classes of functions with mixed smoothness. We use technique, based on a combination of results from hyperbolic cross approximation, which were obtained in 1980s – 1990s, and recent results on greedy approximation to obtain sharp estimates for best m -term approximation with respect to the trigonometric system. We give some observations on numerical integration and approximate recovery of functions with mixed smoothness. We prove lower bounds, which show that one cannot improve accuracy of sparse grids methods with $\asymp 2^n n^{d-1}$ points in the grid by adding 2^n arbitrary points. In case of numerical integration these lower bounds provide best known lower bounds for optimal cubature formulas and for sparse grids based cubature formulas.

1 Introduction

Sparse approximation with respect to dictionaries is a very important topic in the high-dimensional approximation. The main motivation for the study

**Math Subject Classifications. primary: 41A65; secondary: 42A10, 46B20.*

[†]University of South Carolina, USA, and Steklov Institute of Mathematics, Russia. Research was supported by NSF grant DMS-1160841

of sparse approximation is that many real world signals can be well approximated by sparse ones. Sparse approximation automatically implies a need for nonlinear approximation, in particular, for greedy approximation. We give a brief description of a sparse approximation problem and present a discussion of the obtained results and their relation to previous work. In a general setting we are working in a Banach space X with a redundant system of elements \mathcal{D} (dictionary \mathcal{D}). There is a solid justification of importance of a Banach space setting in numerical analysis in general and in sparse approximation in particular (see, for instance, [28], Preface). Let X be a real Banach space with norm $\|\cdot\| := \|\cdot\|_X$. We say that a set of elements (functions) \mathcal{D} from X is a dictionary if each $g \in \mathcal{D}$ has norm one ($\|g\| = 1$), and the closure of $\text{span } \mathcal{D}$ is X . A symmetrized dictionary is $\mathcal{D}^\pm := \{\pm g : g \in \mathcal{D}\}$. For a nonzero element $g \in X$ we let F_g denote a norming (peak) functional for g :

$$\|F_g\|_{X^*} = 1, \quad F_g(g) = \|g\|_X.$$

The existence of such a functional is guaranteed by the Hahn-Banach theorem.

An element (function, signal) $s \in X$ is said to be m -sparse with respect to \mathcal{D} if it has a representation $s = \sum_{i=1}^m c_i g_i$, $g_i \in \mathcal{D}$, $i = 1, \dots, m$. The set of all m -sparse elements is denoted by $\Sigma_m(\mathcal{D})$. For a given element f we introduce the error of best m -term approximation

$$\sigma_m(f, \mathcal{D})_X := \inf_{s \in \Sigma_m(\mathcal{D})} \|f - s\|_X.$$

For a function class \mathbf{W} define

$$\sigma_m(\mathbf{W}, \mathcal{D})_X := \sup_{f \in \mathbf{W}} \sigma_m(f, \mathcal{D})_X.$$

Let $t \in (0, 1]$ be a given nonnegative number. We define (see [25]) the Weak Chebyshev Greedy Algorithm (WCGA) that is a generalization for Banach spaces of Weak Orthogonal Greedy Algorithm defined and studied in [24] (see also [28]).

Weak Chebyshev Greedy Algorithm (WCGA). We define $f_0 := f_0^{c,t} := f$. Then for each $m \geq 1$ we inductively define

- 1). $\varphi_m := \varphi_m^{c,t} \in \mathcal{D}$ is any element satisfying

$$|F_{f_{m-1}}(\varphi_m)| \geq t \sup_{g \in \mathcal{D}} |F_{f_{m-1}}(g)|.$$

2). Define

$$\Phi_m := \Phi_m^t := \text{span}\{\varphi_j\}_{j=1}^m,$$

and define $G_m := G_m^{c,t}$ to be the best approximant to f from Φ_m .

3). Denote

$$f_m := f_m^{c,t} := f - G_m.$$

We demonstrated in the paper [30] that the Weak Chebyshev Greedy Algorithm (WCGA) is very good for m -term approximation with respect to a special class of dictionaries, in particular, for the trigonometric system. The trigonometric system is a classical system that is known to be difficult to study. In [30] we study among other problems the problem of nonlinear sparse approximation with respect to it. Let \mathcal{RT} denote the real trigonometric system $1, \sin 2\pi x, \cos 2\pi x, \dots$ on $[0, 1]$ and let \mathcal{RT}_p to be its version normalized in $L_p([0, 1])$. Denote $\mathcal{RT}_p^d := \mathcal{RT}_p \times \dots \times \mathcal{RT}_p$ the d -variate trigonometric system. We need to consider the real trigonometric system because the algorithm WCGA is well studied for the real Banach space. We proved in [30] the following Lebesgue-type inequality for the WCGA.

Theorem 1.1. *Let \mathcal{D} be the normalized in L_p , $2 \leq p < \infty$, real d -variate trigonometric system. Then for any $f \in L_p$ the WCGA with weakness parameter t gives*

$$\|f_{C(t,p,d)m \ln(m+1)}\|_p \leq C\sigma_m(f, \mathcal{D})_p. \quad (1.1)$$

The above Lebesgue-type inequality guarantees that the WCGA works very well for each individual function f . As a complement to this inequality we would like to obtain results, which relate rate of decay of $\sigma_m(f, \mathcal{T}^d)_p$ to some smoothness type properties of f . It is the main goal of this paper. We measure smoothness in terms of mixed derivative and mixed difference. We note that the function classes with bounded mixed derivative are not only interesting and challenging object for approximation theory but they are important in numerical computations. Griebel and his group use approximation methods designed for these classes in elliptic variational problems. Recent work of Yserentant on new regularity models for the Schrödinger equation shows that the eigenfunctions of the electronic Schrödinger operator have a certain mixed smoothness similar to the bounded mixed derivative. This makes approximation techniques developed for classes of functions with bounded mixed derivative a proper choice for the numerical treatment of the Schrödinger equation.

Sparse trigonometric approximation of periodic functions began by the paper of S.B. Stechkin [13], who used it in the criterion for absolute convergence of trigonometric series. R.S. Ismagilov [8] found nontrivial estimates for m -term approximation of functions with singularities of the type $|x|$ and gave interesting and important applications to the widths of Sobolev classes. He used a deterministic method based on number theoretical constructions. His method was developed by V.E. Maiorov [10], who used a method based on Gaussian sums. Further strong results were obtained in [3] with the help of a nonconstructive result from finite dimensional Banach spaces due to E.D. Gluskin [7]. Other powerful nonconstructive method, which is based on a probabilistic argument, was used by Y. Makovoz [11] and by E.S. Belinskii [2]. Different methods were created in [14], [9], [23], [29] for proving lower bounds for function classes. It was discovered in [4] and [27] that greedy algorithms can be used for constructive m -term approximation with respect to the trigonometric system. We demonstrate in this paper how greedy algorithms can be used to prove optimal or best known upper bounds for m -term approximation of classes of function with mixed smoothness. It is a simple and powerful method of proving upper bounds. However, we do not know how to use it for small smoothness. The reader can find a detailed study of m -term approximation of classes of function with mixed smoothness, including small smoothness, in the paper [12] by A.S. Romanyuk. We note that in the case $2 < p < \infty$ the upper bounds in [12] are not constructive.

We begin with some notation. Let $\mathbf{s} = (s_1, \dots, s_d)$ be a vector whose coordinates are nonnegative integers

$$\rho(\mathbf{s}) := \{\mathbf{k} \in \mathbb{Z}^d : [2^{s_j-1}] \leq |k_j| < 2^{s_j}, \quad j = 1, \dots, d\},$$

$$Q_n := \cup_{\|\mathbf{s}\|_1 \leq n} \rho(\mathbf{s}) \quad - \quad \text{a step hyperbolic cross},$$

$$\Gamma(N) := \{\mathbf{k} \in \mathbb{Z}^d : \prod_{j=1}^d \max(|k_j|, 1) \leq N\} \quad - \quad \text{a hyperbolic cross}.$$

For $f \in L_1(\mathbb{T}^d)$

$$\delta_{\mathbf{s}}(f, \mathbf{x}) := \sum_{\mathbf{k} \in \rho(\mathbf{s})} \hat{f}(\mathbf{k}) e^{i(\mathbf{k}, \mathbf{x})}.$$

Let G be a finite set of points in \mathbb{Z}^d , we denote

$$\mathcal{T}(G) := \left\{ t : t(\mathbf{x}) = \sum_{\mathbf{k} \in G} c_{\mathbf{k}} e^{i(\mathbf{k}, \mathbf{x})} \right\}.$$

For the sake of simplicity we shall write $\mathcal{T}(\Gamma(N)) = \mathcal{T}(N)$.

We study some approximation problems for classes of functions with mixed smoothness. We define these classes momentarily. We will begin with the case of univariate periodic functions. Let for $r > 0$

$$F_r(x) := 1 + 2 \sum_{k=1}^{\infty} k^{-r} \cos(kx - r\pi/2) \quad (1.2)$$

and

$$W_p^r := \{f : f = \varphi * F_r, \quad \|\varphi\|_p \leq 1\}. \quad (1.3)$$

It is well known that for $r > 1/p$ the class W_p^r is embedded into the space of continuous functions $C(\mathbb{T})$. In a particular case of W_1^1 we also have embedding into $C(\mathbb{T})$.

In the multivariate case for $\mathbf{x} = (x_1, \dots, x_d)$ denote

$$F_r(\mathbf{x}) := \prod_{j=1}^d F_r(x_j)$$

and

$$\mathbf{W}_p^r := \{f : f = \varphi * F_r, \quad \|\varphi\|_p \leq 1\}.$$

For $f \in \mathbf{W}_p^r$ we will denote $f^{(r)} := \varphi$ where φ is such that $f = \varphi * F_r$.

The main results of Section 2 are the following two theorems. We use the notation $\beta := \beta(q, p) := 1/q - 1/p$ and $\eta := \eta(q) := 1/q - 1/2$. In the case of trigonometric system \mathcal{T}^d we drop it from the notation:

$$\sigma_m(\mathbf{W})_p := \sigma_m(\mathbf{W}, \mathcal{T}^d)_p.$$

Theorem 1.2. *We have*

$$\sigma_m(\mathbf{W}_q^r)_p \asymp \begin{cases} m^{-r+\beta}(\log m)^{(d-1)(r-2\beta)}, & 1 < q \leq p \leq 2, \quad r > 2\beta, \\ m^{-r+\eta}(\log m)^{(d-1)(r-2\eta)}, & 1 < q \leq 2 \leq p < \infty, \quad r > 1/q, \\ m^{-r}(\log m)^{r(d-1)}, & 2 \leq q \leq p < \infty, \quad r > 1/2. \end{cases}$$

Theorem 1.3. *We have*

$$\sigma_m(\mathbf{W}_q^r)_\infty \ll \begin{cases} m^{-r+\eta}(\log m)^{(d-1)(r-2\eta)+1/2}, & 1 < q \leq 2, \quad r > 1/q, \\ m^{-r}(\log m)^{r(d-1)+1/2}, & 2 \leq q < \infty, \quad r > 1/2. \end{cases}$$

The case $1 < q \leq p \leq 2$ in Theorem 1.2, which corresponds to the first line, was proved in [14] (see also [16], Ch.4). The proofs from [14] and [16] are constructive. In Section 2 we concentrate on the case $p \geq 2$. We use recently developed techniques on greedy approximation in Banach spaces to prove Theorems 1.2 and 1.3. It is important that greedy approximation allows us not only to prove the above theorems but also to provide a constructive way for building the corresponding m -term approximants. We give a precise formulation.

Theorem 1.4. *For $p \in (1, \infty)$ and $\mu > 0$ there exist constructive methods $A_m(f, p, \mu)$, which provide for $f \in \mathbf{W}_q^r$ an m -term approximation such that*

$$\|f - A_m(f, p, \mu)\|_p \ll \begin{cases} m^{-r+\beta}(\log m)^{(d-1)(r-2\beta)}, & 1 < q \leq p \leq 2, \quad r > 2\beta + \mu, \\ m^{-r+\eta}(\log m)^{(d-1)(r-2\eta)}, & 1 < q \leq 2 \leq p < \infty, \quad r > 1/q + \mu, \\ m^{-r}(\log m)^{r(d-1)}, & 2 \leq q \leq p < \infty, \quad r > 1/2 + \mu. \end{cases}$$

Similar modification of Theorem 1.3 holds for $p = \infty$. We do not have matching lower bounds for the upper bounds in Theorem 1.3 in the case of approximation in the uniform norm L_∞ . In Section 3 we use known results on the entropy numbers to prove one lower bound in the case of functions of two variables. We note that it is of interest for small smoothness: $r < 1/2$.

As a direct corollary of Theorems 1.1 and 1.2 we obtain the following result.

Theorem 1.5. *Let $p \in [2, \infty)$. Apply the WCGA with weakness parameter $t \in (0, 1]$ to $f \in L_p$ with respect to the real trigonometric system \mathcal{RT}_p^d . If $f \in \mathbf{W}_q^r$, then we have*

$$\|f_m\|_p \ll \begin{cases} m^{-r+\eta}(\log m)^{(d-1)(r-2\eta)+r-\eta}, & 1 < q \leq 2, \quad r > 1/q, \\ m^{-r}(\log m)^{rd}, & 2 \leq q < \infty, \quad r > 1/2. \end{cases}$$

In Sections 4 and 5 we give some observations on numerical integration and approximate recovery of functions with mixed smoothness. We prove there some lower bounds, which show that we cannot improve accuracy of sparse grids methods with $\asymp 2^n n^{d-1}$ points in the grid by adding 2^n arbitrary points. In case of numerical integration these lower bounds provide best known lower bounds for optimal cubature formulas and for sparse grids based cubature formulas.

Our technique is based on a combination of results from the hyperbolic cross approximation, which were obtained in 1980s – 1990s, and recent results on greedy approximation. We formulate some known results from the hyperbolic cross approximation theory, which will be used in our analysis. We begin with the problem of estimating $\|f\|_p$ in terms of the array $\{\|\delta_{\mathbf{s}}(f)\|_q\}$. Here and below p and q are scalars such that $1 \leq q, p \leq \infty$. Let an array $\varepsilon = \{\varepsilon_{\mathbf{s}}\}$ be given, where $\varepsilon_{\mathbf{s}} \geq 0$, $\mathbf{s} = (s_1, \dots, s_d)$, and s_j are nonnegative integers, $j = 1, \dots, d$. We denote by $G(\varepsilon, q)$ and $F(\varepsilon, q)$ the following sets of functions ($1 \leq q \leq \infty$):

$$G(\varepsilon, q) := \{f \in L_q : \|\delta_{\mathbf{s}}(f)\|_q \leq \varepsilon_{\mathbf{s}} \quad \text{for all } \mathbf{s}\},$$

$$F(\varepsilon, q) := \{f \in L_q : \|\delta_{\mathbf{s}}(f)\|_q \geq \varepsilon_{\mathbf{s}} \quad \text{for all } \mathbf{s}\}.$$

The following theorem is from [16], p.29. For the special case $q = 2$, which will be used in this paper, see [14] and [16], p.86.

Theorem 1.6. *The following relations hold:*

$$\sup_{f \in G(\varepsilon, q)} \|f\|_p \asymp \left(\sum_{\mathbf{s}} \varepsilon_{\mathbf{s}}^p 2^{\|\mathbf{s}\|_1(p/q-1)} \right)^{1/p}, \quad 1 \leq q < p < \infty; \quad (1.4)$$

$$\inf_{f \in F(\varepsilon, q)} \|f\|_p \asymp \left(\sum_{\mathbf{s}} \varepsilon_{\mathbf{s}}^p 2^{\|\mathbf{s}\|_1(p/q-1)} \right)^{1/p}, \quad 1 < p < q \leq \infty, \quad (1.5)$$

with constants independent of ε .

We will need a corollary of Theorem 1.6 (see [16], Ch.1, Theorem 2.2), which we formulate as a theorem.

Theorem 1.7. *Let $1 < q \leq 2$. For any $t \in \mathcal{T}(N)$ we have*

$$\|t\|_A := \sum_{\mathbf{k}} |\hat{t}(\mathbf{k})| \leq C(q, d) N^{1/q} (\log N)^{(d-1)(1-1/q)} \|t\|_q.$$

Let

$$\Pi(\mathbf{N}, d) := \{(a_1, \dots, a_d) \in \mathbb{R}^d : |a_j| \leq N_j, j = 1, \dots, d\},$$

where N_j are nonnegative integers and $\mathbf{N} := (N_1, \dots, N_d)$. We denote

$$\mathcal{T}(\mathbf{N}, d) := \{t : t = \sum_{\mathbf{k} \in \Pi(\mathbf{N}, d)} c_{\mathbf{k}} e^{i(\mathbf{k}, \mathbf{x})}\}.$$

Then

$$\dim \mathcal{T}(\mathbf{N}, d) = \prod_{j=1}^d (2N_j + 1) =: \vartheta(\mathbf{N}).$$

The following theorem is from [22], Ch.2, Theorem 1.1 (see, also, [17]).

Theorem 1.8. *Let $\varepsilon > 0$ and a subspace $\Psi \subset \mathcal{T}(\mathbf{N}, d)$ be such that $\dim \Psi \geq \varepsilon \dim \mathcal{T}(\mathbf{N}, d)$. Then there is a $t \in \Psi$ such that*

$$\|t\|_\infty = 1, \quad \|t\|_2 \geq C(\varepsilon, d) > 0.$$

2 Sparse approximation

For a Banach space X we define the modulus of smoothness

$$\rho(u) := \sup_{\|x\|=\|y\|=1} \left(\frac{1}{2} (\|x + uy\| + \|x - uy\|) - 1 \right).$$

The uniformly smooth Banach space is the one with the property

$$\lim_{u \rightarrow 0} \rho(u)/u = 0.$$

It is well known (see for instance [6], Lemma B.1) that in the case $X = L_p$, $1 \leq p < \infty$, we have

$$\rho(u) \leq \begin{cases} u^p/p & \text{if } 1 \leq p \leq 2, \\ (p-1)u^2/2 & \text{if } 2 \leq p < \infty. \end{cases} \quad (2.1)$$

Denote by $A_1(\mathcal{D}) := A_1(\mathcal{D}, X)$ the closure in X of the convex hull of \mathcal{D} . The following theorem from [25] gives the rate of convergence of the WCGA for f in $A_1(\mathcal{D}^\pm)$.

Theorem 2.1. *Let X be a uniformly smooth Banach space with the modulus of smoothness $\rho(u) \leq \gamma u^q$, $1 < q \leq 2$. Then for $t \in (0, 1]$ we have for any $f \in A_1(\mathcal{D}^\pm)$ that*

$$\|f - G_m^{c,t}(f, \mathcal{D})\| \leq C(q, \gamma)(1 + mt^p)^{-1/p}, \quad p := \frac{q}{q-1},$$

with a constant $C(q, \gamma)$ which may depend only on q and γ .

Remark 2.1. *It follows from the proof of Theorem 2.1 that $C(q, \gamma) \leq C(q)\gamma^{1/q}$.*

We proceed to the Incremental Greedy Algorithm (see [27] and [28], Chapter 6). Let $\epsilon = \{\epsilon_n\}_{n=1}^\infty$, $\epsilon_n > 0$, $n = 1, 2, \dots$.

Incremental Algorithm with schedule ϵ (IA(ϵ)). Denote $f_0^{i,\epsilon} := f$ and $G_0^{i,\epsilon} := 0$. Then, for each $m \geq 1$ we have the following inductive definition.

(1) $\varphi_m^{i,\epsilon} \in \mathcal{D}$ is any element satisfying

$$F_{f_{m-1}^{i,\epsilon}}(\varphi_m^{i,\epsilon} - f) \geq -\epsilon_m.$$

(2) Define

$$G_m^{i,\epsilon} := (1 - 1/m)G_{m-1}^{i,\epsilon} + \varphi_m^{i,\epsilon}/m.$$

(3) Let

$$f_m^{i,\epsilon} := f - G_m^{i,\epsilon}.$$

Theorem 2.2. *Let X be a uniformly smooth Banach space with modulus of smoothness $\rho(u) \leq \gamma u^q$, $1 < q \leq 2$. Define*

$$\epsilon_n := v\gamma^{1/q}n^{-1/p}, \quad p = \frac{q}{q-1}, \quad n = 1, 2, \dots$$

Then, for any $f \in A_1(\mathcal{D})$, we have

$$\|f_m^{i,\epsilon}\| \leq C(v)\gamma^{1/q}m^{-1/p}, \quad m = 1, 2, \dots$$

In [27] we demonstrated the power of the WCGA in classical areas of harmonic analysis. The problem concerns the trigonometric m -term approximation in the uniform norm. The first result that indicated an advantage of m -term approximation with respect to the real trigonometric system \mathcal{RT} over approximation by trigonometric polynomials of order m is due to Ismagilov [8]

$$\sigma_m(|\sin 2\pi x|, \mathcal{RT})_\infty \leq C_\epsilon m^{-6/5+\epsilon}, \quad \text{for any } \epsilon > 0. \quad (2.2)$$

Maiorov [10] improved the estimate (2.2):

$$\sigma_m(|\sin 2\pi x|, \mathcal{RT})_\infty \asymp m^{-3/2}. \quad (2.3)$$

Both R.S. Ismagilov [8] and V.E. Maiorov [10] used constructive methods to get their estimates (2.2) and (2.3). V.E. Maiorov [10] applied a number theoretical method based on Gaussian sums. The key point of that technique

can be formulated in terms of best m -term approximation of trigonometric polynomials. Let $\mathcal{RT}(N)$ be the subspace of real trigonometric polynomials of order N . Using the Gaussian sums one can prove (constructively) the estimate

$$\sigma_m(t, \mathcal{RT})_\infty \leq CN^{3/2}m^{-1}\|t\|_1, \quad t \in \mathcal{RT}(N). \quad (2.4)$$

Denote

$$\|a_0 + \sum_{k=1}^N (a_k \cos k2\pi x + b_k \sin k2\pi x)\|_A := |a_0| + \sum_{k=1}^N (|a_k| + |b_k|).$$

We note that by simple inequality

$$\|t\|_A \leq CN\|t\|_1, \quad t \in \mathcal{RT}(N),$$

the estimate (2.4) follows from the estimate

$$\sigma_m(t, \mathcal{RT})_\infty \leq C(N^{1/2}/m)\|t\|_A, \quad t \in \mathcal{RT}(N). \quad (2.5)$$

Thus, (2.5) is stronger than (2.4). The following estimate was proved in [3]

$$\sigma_m(t, \mathcal{RT})_\infty \leq Cm^{-1/2}(\ln(1 + N/m))^{1/2}\|t\|_A, \quad t \in \mathcal{RT}(N). \quad (2.6)$$

In a way (2.6) is much stronger than (2.5) and (2.4). The proof of (2.6) from [3] is not constructive. The estimate (2.6) has been proved in [3] with the help of a nonconstructive theorem of Gluskin [7]. E.S. Belinskii [2] used a probabilistic method to prove the following inequality: for $2 \leq p < \infty$

$$\sigma_m(t, \mathcal{RT})_\infty \leq C(N/m)^{1/p}(\ln(1 + N/m))^{1/p}\|t\|_p, \quad t \in \mathcal{RT}(N).$$

His proof is nonconstructive as well. In [27] we gave a constructive proof of (2.6). The key ingredient of that proof is the WCGA. In the paper [4] we already pointed out that the WCGA provides a constructive proof of the estimate

$$\sigma_m(f, \mathcal{RT})_p \leq C(p)m^{-1/2}\|f\|_A, \quad p \in [2, \infty). \quad (2.7)$$

The known proofs (before [4]) of (2.7) were nonconstructive (see discussion in [4], Section 5). Thus, the WCGA provides a way of building a good m -term approximant. We formulate here a result from [27].

Theorem 2.3. *There exists a constructive method $A(N, m)$ such that for any $t \in \mathcal{RT}(N)$ it provides an m -term trigonometric polynomial $A(N, m)(t)$ with the following approximation property*

$$\|t - A(N, m)(t)\|_\infty \leq C m^{-1/2} (\ln(1 + N/m))^{1/2} \|t\|_A$$

with an absolute constant C .

However, the step 2) of the WCGA makes it difficult to control the coefficients of the approximant – they are obtained through the Chebyshev projection of f onto Φ_m . This motivates us to consider the $\text{IA}(\epsilon)$ which gives explicit coefficients of the approximant. An advantage of the $\text{IA}(\epsilon)$ over other greedy-type algorithms is that the $\text{IA}(\epsilon)$ gives precise control of the coefficients of the approximant. For all approximants $G_m^{i, \epsilon}$ we have the property $\|G_m^{i, \epsilon}\|_A = 1$. Moreover, we know that all nonzero coefficients of the approximant have the form a/m where a is a natural number. We prove the following result.

Theorem 2.4. *For any $t \in \mathcal{RT}(\mathbf{N}, d)$ the $\text{IA}(\epsilon)$ applied to $f := t/\|t\|_A$ provides after m iterations an m -term trigonometric polynomial $G_m(t) := G_m^{i, \epsilon}(f)\|t\|_A$ with the following approximation property*

$$\|t - G_m(t)\|_\infty \leq C(d)(\bar{m})^{-1/2} (\ln \vartheta(\mathbf{N}))^{1/2} \|t\|_A,$$

$$\bar{m} := \max(1, m), \quad \|G_m(t)\|_A = \|t\|_A,$$

with a constant $C(d)$, which may depend only on d .

Proof. It is clear that it is sufficient to prove Theorem 2.4 for $t \in \mathcal{RT}(\mathbf{N}, d)$ with $\|t\|_A = 1$. Then $t \in (A_1(\mathcal{RT}(\mathbf{N}, d) \cap \mathcal{RT}^d)^\pm, L_p)$ for all $p \in [2, \infty)$. Now, applying Theorem 2.2 with $X = L_p$ and \mathcal{D}^\pm , where $\mathcal{D} := \{\varphi_1, \varphi_2, \dots, \varphi_n\}$, $n = \vartheta(\mathbf{N})$, being the real trigonometric system

$$\varphi_l := \prod_{j \in E} \cos k_j x_j \prod_{j \in [1, d] \setminus E} \sin k_j x_j,$$

we obtain that

$$\|t - \sum_{j \in \Lambda} \frac{a_j}{m} \varphi_j\|_p \leq C \gamma^{1/2} m^{-1/2}, \quad \sum_{j \in \Lambda} |a_j| = m, \quad (2.8)$$

where $\sum_{j \in \Lambda} \frac{a_j}{m} \varphi_j$ is the $G_m^{i, \epsilon}(t)$. By (2.1) we find $\gamma \leq p/2$. Next, by the Nikol'skii inequality we get from (2.8)

$$\|t - \sum_{j \in \Lambda} \frac{a_j}{m} \varphi_j\|_\infty \leq C(d) N^{1/p} \|t - \sum_{j \in \Lambda} \frac{a_j}{m} \varphi_j\|_p \leq C(d) p^{1/2} N^{1/p} m^{-1/2}.$$

Choosing $p \asymp \ln N$ we obtain the desired in Theorem 2.4 bound. \square

We point out that the above proof of Theorem 2.4 gives the following statement.

Theorem 2.5. *Let $2 \leq p < \infty$. For any $t \in \mathcal{RT}(\mathbf{N}, d)$ the $IA(\epsilon)$ applied to $f := t/\|t\|_A$ provides after m iterations an m -term trigonometric polynomial $G_m(t) := G_m^{i, \epsilon}(f)\|t\|_A$ with the following approximation property*

$$\|t - G_m(t)\|_p \leq C(d)(\bar{m})^{-1/2} p^{1/2} \|t\|_A, \quad \bar{m} := \max(1, m), \quad \|G_m(t)\|_A = \|t\|_A,$$

with a constant $C(d)$, which may depend only on d .

We note that the implementation of the $IA(\epsilon)$ depends on the dictionary and the ambient space X . For example, for $d = 1$ the $IA(\epsilon)$ from Theorem 2.4 acts with respect to the real trigonometric system $1, \cos 2\pi x, \sin 2\pi x, \dots, \cos N2\pi x, \sin N2\pi x$ in the space $X = L_p$ with $p \asymp \ln N$.

The above Theorems 2.4 and 2.5 are formulated for m -term approximation with respect to the real trigonometric system because the general Theorem 2.2 is proved for real Banach spaces. Clearly, as a corollary of Theorems 2.4 and 2.5 we obtain the corresponding results for the complex trigonometric system $\mathcal{T}^d := \{e^{i(\mathbf{k}, \mathbf{x})}\}_{\mathbf{k} \in \mathbb{Z}^d}$.

Theorem 2.6. *There exist constructive greedy-type approximation methods $G_m^p(\cdot)$, which provide m -term polynomials with respect to \mathcal{T}^d with the following properties: for $2 \leq p < \infty$*

$$\|f - G_m^p(f)\|_p \leq C_1(d)(\bar{m})^{-1/2} p^{1/2} \|f\|_A, \quad \|G_m^p(f)\|_A \leq C_2(d) \|f\|_A, \quad (2.9)$$

and for $p = \infty$, $f \in \mathcal{T}(\mathbf{N}, d)$

$$\|f - G_m^\infty(f)\|_\infty \leq C_3(d)(\bar{m})^{-1/2} (\ln \vartheta(\mathbf{N}))^{1/2} \|f\|_A, \quad \|G_m^\infty(f)\|_A \leq C_4(d) \|f\|_A. \quad (2.10)$$

We now apply Theorem 2.6 for m -term approximation of functions with mixed smoothness. The following theorem was proved in [14] (see also [16], Ch.4). The proofs from [14] and [16] are constructive. We use the following notation $\beta := \beta(q, p) := 1/q - 1/p$ and $\eta := \eta(q) := 1/q - 1/2$.

Theorem 2.7. *Let $1 < q \leq p \leq 2$, $r > 2\beta$. Then*

$$\sigma_m(\mathbf{W}_q^r)_p \asymp m^{-r+\beta}(\log m)^{(d-1)(r-2\beta)}.$$

First, we extend Theorem 2.7 to the case $1 < q \leq p < \infty$.

Theorem 2.8. *One has*

$$\sigma_m(\mathbf{W}_q^r)_p \asymp \begin{cases} m^{-r+\beta}(\log m)^{(d-1)(r-2\beta)}, & 1 < q \leq p \leq 2, \quad r > 2\beta, \\ m^{-r+\eta}(\log m)^{(d-1)(r-2\eta)}, & 1 < q \leq 2 \leq p < \infty, \quad r > 1/q, \\ m^{-r}(\log m)^{r(d-1)}, & 2 \leq q \leq p < \infty, \quad r > 1/2. \end{cases}$$

Proof. The case $p \leq 2$, which corresponds to the first line, follows from Theorem 2.7. We note that in the case $p > 2$ Theorem 2.8 is proved in [23]. However, the proof there is not constructive – it uses a nonconstructive result from [3]. We provide a constructive proof, which is based on greedy algorithms. Also, this proof works under weaker conditions on r : $r > 1/q$ instead of $r > 1/q + \eta$ for $1 < q \leq 2$. The following lemma plays the key role in the proof.

Lemma 2.1. *Define for $f \in L_1$*

$$f_l := \sum_{\|\mathbf{s}\|_1=l} \delta_{\mathbf{s}}(f), \quad l \in \mathbb{N}_0, \quad \mathbb{N}_0 := \mathbb{N} \cup \{0\}.$$

Consider the class

$$\mathbf{W}_A^{a,b} := \{f : \|f_l\|_A \leq 2^{-al} l^{(d-1)b}\}.$$

Then for $2 \leq p \leq \infty$ and $0 < \mu < a$ there is a constructive method $A_m(\cdot, p, \mu)$ based on greedy algorithms, which provides the bound for $f \in \mathbf{W}_A^{a,b}$

$$\|f - A_m(f, p, \mu)\|_p \ll m^{-a-1/2}(\log m)^{(d-1)(a+b)}, \quad 2 \leq p < \infty, \quad (2.11)$$

$$\|f - A_m(f, p, \mu)\|_\infty \ll m^{-a-1/2}(\log m)^{(d-1)(a+b)+1/2}. \quad (2.12)$$

Proof. We prove the lemma for $m \asymp 2^n n^{d-1}$, $n \in \mathbb{N}$. Let $f \in \mathbf{W}_A^{a,b}$. We approximate f_l in L_p . By Theorem 2.6 we obtain for $p \in [2, \infty)$

$$\|f_l - G_{m_l}^p(f_l)\|_p \ll (\bar{m}_l)^{-1/2} \|f_l\|_A \ll (\bar{m}_l)^{-1/2} 2^{-al} l^{(d-1)b}. \quad (2.13)$$

We take $\mu \in (0, a)$ and specify

$$m_l := [2^{n-\mu(l-n)} l^{d-1}], \quad l = n, n+1, \dots$$

In addition we include in the approximant

$$S_n(f) := \sum_{\|\mathbf{s}\|_1 \leq n} \delta_{\mathbf{s}}(f).$$

Define

$$A_m(f, p, \mu) := S_n(f) + \sum_{l > n} G_{m_l}^p(f_l).$$

Then, we have built an m -term approximant of f with

$$m \ll 2^n n^{d-1} + \sum_{l \geq n} m_l \ll 2^n n^{d-1}.$$

The error of this approximation in L_p is bounded from above by

$$\begin{aligned} \|f - A_m(f, p, \mu)\|_p &\leq \sum_{l \geq n} \|f_l - G_{m_l}^p(f_l)\|_p \ll \sum_{l \geq n} (\bar{m}_l)^{-1/2} 2^{-al} l^{(d-1)b} \\ &\ll \sum_{l \geq n} 2^{-1/2(n-\mu(l-n))} l^{-(d-1)/2} 2^{-al} l^{(d-1)b} \ll 2^{-n(a+1/2)} n^{(d-1)(b-1/2)}. \end{aligned}$$

This completes the proof of lemma in the case $2 \leq p < \infty$.

Let us discuss the case $p = \infty$. The proof repeats the proof in the above case $p < \infty$ with the following change. Instead of using (2.9) for estimating an m_l -term approximation of f_l in L_p we use (2.10) to estimate an m_l -term approximation of f_l in L_∞ . Then bound (2.13) is replaced by

$$\|f_l - G_{m_l}^\infty(f_l)\|_\infty \ll (\bar{m}_l)^{-1/2} (\ln 2^l)^{1/2} \|f_l\|_A \ll (\bar{m}_l)^{-1/2} l^{1/2} 2^{-al} l^{(d-1)b}. \quad (2.14)$$

The extra factor $l^{1/2}$ in (2.14) gives an extra factor $(\log m)^{1/2}$ in (2.12). \square

We now complete the proof of Theorem 2.8. First, consider the case $1 < q \leq 2 \leq p < \infty$. It is well known (see, for instance, [16], p.34, Theorem 2.1) that for $f \in \mathbf{W}_q^r$ one has

$$\|f_l\|_q \ll 2^{-lr}. \quad (2.15)$$

Theorem 1.7 implies

$$\|f_l\|_A \ll 2^{-(r-1/q)l} l^{(d-1)(1-1/q)}.$$

Therefore, it is sufficient to use Lemma 2.1 with $a = r - 1/q$ and $b = 1 - 1/q$ to obtain the upper bounds.

Second, the upper bounds in the case $2 \leq q \leq p < \infty$ follow from the above case $1 < q \leq 2 \leq p < \infty$ with $q = 2$. The lower bounds follow from Theorem 2.7 with $p = 2$. The lower bounds in the case $2 \leq q \leq p < \infty$ follow from known results for the case $1 < p \leq q < \infty$ in [9] (see Theorem 2.17 below). \square

Let us discuss the case $p = \infty$. In the same way as Theorem 2.8 was derived from (2.11) of Lemma 2.1 the following upper bounds in case $p = \infty$ are derived from (2.12) of Lemma 2.1.

Theorem 2.9. *We have*

$$\sigma_m(\mathbf{W}_q^r)_\infty \ll \begin{cases} m^{-r+\eta}(\log m)^{(d-1)(r-2\eta)+1/2}, & 1 < q \leq 2, \quad r > 1/q, \\ m^{-r}(\log m)^{r(d-1)+1/2}, & 2 \leq q < \infty, \quad r > 1/2. \end{cases}$$

The upper bounds are provided by a constructive method $A_m(\cdot, \infty, \mu)$ based on greedy algorithms.

Consider the case $\sigma_m(\mathbf{W}_{1,\alpha}^r)_p$, which is not covered by Theorems 2.7 and 2.8. The function $F_r(\mathbf{x})$ belongs to the closure in L_p of $\mathbf{W}_{1,\alpha}^r$, $r > 1 - 1/p$, and, therefore, on the one hand

$$\sigma_m(\mathbf{W}_{1,\alpha}^r)_p \geq \sigma_m(F_r(\mathbf{x}))_p.$$

On the other hand, it follows from the definition of $\mathbf{W}_{1,\alpha}^r$ that for any $f \in \mathbf{W}_{1,\alpha}^r$

$$\sigma_m(f)_p \leq \sigma_m(F_r(\mathbf{x}))_p.$$

Thus,

$$\sigma_m(\mathbf{W}_{1,\alpha}^r)_p = \sigma_m(F_r(\mathbf{x}))_p. \quad (2.16)$$

We now prove some results on $\sigma_m(F_r(\mathbf{x}))_p$.

Theorem 2.10. *We have*

$$\sigma_m(F_r(\mathbf{x}))_p \asymp \begin{cases} m^{-r+1-1/p}(\log m)^{(d-1)(r-1+2/p)}, & 1 < p \leq 2, \quad r > 1 - 1/p, \\ m^{-r+1/2}(\log m)^{r(d-1)}, & 2 \leq p < \infty, \quad r > 1. \end{cases}$$

The upper bounds are provided by a constructive method $A_m(\cdot, p, \mu)$ based on greedy algorithms.

Proof. We begin with the case $1 < p \leq 2$. The following error bound for approximation by the hyperbolic cross polynomials is known (see, for instance, [16], p.38)

$$E_{Q_n}(F_r)_p := \inf_{t \in \mathcal{T}(Q_n)} \|F_r - t\|_p \ll 2^{-n(r-1+1/p)} n^{(d-1)/p}. \quad (2.17)$$

Taking into account that $|Q_n| \asymp 2^n n^{d-1}$ we obtain from (2.17) the required upper bound in the case $1 < p \leq 2$. Thus, it remains to prove the matching lower bound in the case $1 < p \leq 2$. Denote

$$\theta_n := \{\mathbf{s} \in \mathbb{N}^d : \|\mathbf{s}\|_1 = n\}, \quad \Delta Q_n := \cup_{\mathbf{s} \in \theta_n} \rho(\mathbf{s}).$$

Let $K_m := \{\mathbf{k}^j\}_{j=1}^m$ be given. Choose n such that it is the minimal to satisfy

$$|\Delta Q_n| \geq 4m.$$

Clearly

$$2^n n^{d-1} \asymp m.$$

Denote

$$\theta'_n := \{\mathbf{s} \in \theta_n : |K_m \cap \rho(\mathbf{s})| \leq |\rho(\mathbf{s})|/2\}.$$

Note that for $\mathbf{s} \in \mathbb{N}^d$ we have $|\rho(\mathbf{s})| = 2^n$. Then

$$(|\theta_n| - |\theta'_n|)2^n/2 \leq m \leq |\theta_n|2^n/4,$$

which implies

$$|\theta'_n| \geq |\theta_n|/2.$$

By Theorem 1.6, (1.5) with $q = 2$, $1 < p < 2$, we obtain for any $t = \sum_{j=1}^m c_j e^{i(\mathbf{k}^j, \mathbf{x})}$

$$\|F_r - t\|_p \gg \left(\sum_{\mathbf{s} \in \theta'_n} \|\delta_{\mathbf{s}}(F_r - t)\|_2^p 2^{n(p/2-1)} \right)^{1/p}$$

$$\gg \left(\sum_{\mathbf{s} \in \theta'_n} 2^{pn(-r+1/2)} 2^{n(p/2-1)} \right)^{1/p} \gg 2^{-n(r-1+1/p)} n^{(d-1)/p}.$$

This gives the required lower bound for $1 < p < 2$. The above argument gives the lower bound in the case $p = 2$ without use of Theorem 1.6 – it is sufficient to use the Parseval identity.

We now proceed to the case $2 \leq p < \infty$. Analysis here is similar to that in the proof of Theorem 2.8. We get for

$$F_r^l := \sum_{\|\mathbf{s}\|_1=l} \delta_{\mathbf{s}}(F_r)$$

$$\|F_r^l\|_A \ll 2^{-lr} 2^l l^{d-1}.$$

The required upper bound follows from Lemma 2.1 with $a = r - 1$ and $b = 1$.

The lower bound follows from the case $p = 2$. \square

In the same way as a modification of the proof of Theorem 2.8 gave Theorem 2.9 the corresponding modification of the argument in the proof of Theorem 2.10 gives the following result.

Theorem 2.11. *We have*

$$\sigma_m(F_r(\mathbf{x}))_\infty \ll m^{-r+1/2} (\log m)^{r(d-1)+1/2}, \quad r > 1.$$

The bounds are provided by a constructive method $A_m(\cdot, \infty, \mu)$ based on greedy algorithms.

We now proceed to classes \mathbf{H}_q^r and $\mathbf{B}_{q,\theta}^r$. Define

$$\|f\|_{\mathbf{H}_q^r} := \sup_{\mathbf{s}} \|\delta_{\mathbf{s}}(f)\|_q 2^{r\|\mathbf{s}\|_1},$$

and for $1 \leq \theta < \infty$ define

$$\|f\|_{\mathbf{B}_{q,\theta}^r} := \left(\sum_{\mathbf{s}} (\|\delta_{\mathbf{s}}(f)\|_q 2^{r\|\mathbf{s}\|_1})^\theta \right)^{1/\theta}.$$

We will write $\mathbf{B}_{q,\infty}^r := \mathbf{H}_q^r$. With a little abuse of notation, denote the corresponding unit ball

$$\mathbf{B}_{q,\theta}^r := \{f : \|f\|_{\mathbf{B}_{q,\theta}^r} \leq 1\}.$$

It will be convenient for us to use the following slight modification of classes $\mathbf{B}_{q,\theta}^r$. Define

$$\|f\|_{\mathbf{H}_{q,\theta}^r} := \sup_n \left(\sum_{\mathbf{s}: \|\mathbf{s}\|_1 = n} (\|\delta_{\mathbf{s}}(f)\|_q 2^{r\|\mathbf{s}\|_1})^\theta \right)^{1/\theta}$$

and

$$\mathbf{H}_{q,\theta}^r := \{f : \|f\|_{\mathbf{H}_{q,\theta}^r} \leq 1\}.$$

The best m -term approximations of classes $\mathbf{B}_{q,\theta}^r$ are studied in detail by A.S. Romanyuk [12]. The following theorem was proved in [14] (see also [16], Ch.4). The proofs from [14] and [16] are constructive.

Theorem 2.12. *Let $1 < q \leq p \leq 2$, $r > \beta$. Then*

$$\sigma_m(\mathbf{H}_q^r)_p \asymp m^{-r+\beta} (\log m)^{(d-1)(r-\beta+1/p)}.$$

The following analog of Theorem 2.8 for classes \mathbf{H}_q^r was proved in [12]. The proof in [12] in the case $p > 2$ is not constructive.

Theorem 2.13. *One has*

$$\sigma_m(\mathbf{H}_q^r)_p \asymp \begin{cases} m^{-r+\beta} (\log m)^{(d-1)(r-\beta+1/p)}, & 1 < q \leq p \leq 2, \quad r > \beta, \\ m^{-r+\eta} (\log m)^{(d-1)(r-1/q+1)}, & 1 < q \leq 2 \leq p < \infty, \quad r > 1/q, \\ m^{-r} (\log m)^{(d-1)(r+1/2)}, & 2 \leq q \leq p < \infty, \quad r > 1/2. \end{cases}$$

Proposition 2.1. *The upper bounds in Theorem 2.13 are provided by a constructive method $A_m(\cdot, p, \mu)$ based on greedy algorithms.*

Proof. The case $p \leq 2$, which corresponds to the first line, follows from Theorem 2.12. We now consider $p \geq 2$. We get from the definition of classes \mathbf{H}_q^r for $1 < q < \infty$:

$$f \in \mathbf{H}_q^r \iff \|\delta_{\mathbf{s}}(f)\|_q \leq 2^{-r\|\mathbf{s}\|_1}.$$

Next,

$$\|\delta_{\mathbf{s}}(f)\|_A \ll 2^{\|\mathbf{s}\|_1/q} \|\delta_{\mathbf{s}}(f)\|_q.$$

Therefore, for $f \in \mathbf{H}_q^r$ we obtain

$$\|f_l\|_A \ll 2^{-(r-1/q)l} l^{d-1}.$$

Applying Lemma 2.1 with $a = r - 1/q$, $b = 1$ in the case $1 < q \leq 2 \leq p < \infty$, we obtain required upper bounds. The upper bounds in the case $2 \leq q \leq p < \infty$ follow from the above case with $q = 2$. The lower bounds in the case $1 < q \leq 2 \leq p < \infty$ follow from Theorem 2.12.

The lower bounds in the case $2 \leq q \leq p < \infty$ follow from known results in [9] (see Theorem 2.18 below). □

In the case $p = \infty$ we have.

Theorem 2.14. *We have*

$$\sigma_m(\mathbf{H}_q^r)_\infty \ll \begin{cases} m^{-r+\eta}(\log m)^{(d-1)(r-1/q+1)+1/2}, & 1 < q \leq 2, \quad r > 1/q, \\ m^{-r}(\log m)^{(r+1/2)(d-1)+1/2}, & 2 \leq q < \infty, \quad r > 1/2. \end{cases}$$

The upper bounds are provided by a constructive method $A_m(\cdot, \infty, \mu)$ based on greedy algorithms.

We now proceed to classes $\mathbf{B}_{q,\theta}^r$. There is the following extension of Theorem 2.13 (see [12]).

Theorem 2.15. *One has*

$$\begin{aligned} \sigma_m(\mathbf{B}_{q,\theta}^r)_p &\asymp \sigma_m(\mathbf{H}_{q,\theta}^r)_p \\ &\asymp \begin{cases} m^{-r+\beta}(\log m)^{(d-1)(r-\beta+1/p-1/\theta)}, & 1 < q \leq p \leq 2, \quad r > \beta, \\ m^{-r+\eta}(\log m)^{(d-1)(r-1/q+1-1/\theta)}, & 1 < q \leq 2 \leq p < \infty, \quad r > 1/q, \\ m^{-r}(\log m)^{(d-1)(r+1/2-1/\theta)}, & 2 \leq q \leq p < \infty, \quad r > 1/2. \end{cases} \end{aligned}$$

Proposition 2.2. *The upper bounds in Theorem 2.15 are provided by a constructive method based on greedy algorithms.*

Proof. In the same way as Theorems 2.7 and 2.12 were proved in [16] one can prove the following result for classes $\mathbf{B}_{q,\theta}^r$ and $\mathbf{H}_{q,\theta}^r$ in the case $1 \leq q \leq p \leq 2$, $p > 1$:

$$\sigma_m(\mathbf{B}_{q,\theta}^r)_p \asymp \sigma_m(\mathbf{H}_{q,\theta}^r)_p \asymp \left(\frac{(\log m)^{d-1}}{m} \right)^{r-\beta} (\log m)^{(d-1)(1/p-1/\theta)}.$$

This proves the statement of Proposition 2.2 for the first relation in Theorem 2.15.

We now consider $p \geq 2$. We get from the definition of classes $\mathbf{H}_{q,\theta}^r$ for $1 < q < \infty$:

$$f \in \mathbf{H}_{q,\theta}^r \iff \left(\sum_{\|\mathbf{s}\|_1=l} \|\delta_{\mathbf{s}}(f)\|_q^\theta \right)^{1/\theta} \leq 2^{-rl}.$$

Next,

$$\begin{aligned} \|f_l\|_A &\leq \sum_{\|\mathbf{s}\|_1=l} \|\delta_{\mathbf{s}}(f)\|_A \ll 2^{l/q} \sum_{\|\mathbf{s}\|_1=l} \|\delta_{\mathbf{s}}(f)\|_q \\ &\ll 2^{l/q} l^{(d-1)(1-1/\theta)} \sum_{\|\mathbf{s}\|_1=l} (\|\delta_{\mathbf{s}}(f)\|_q^\theta)^{1/\theta} \ll 2^{-l(r-1/q)} l^{(d-1)(1-1/\theta)}. \end{aligned}$$

Therefore, for $f \in \mathbf{H}_{q,\theta}^r$ we obtain

$$\|f_l\|_A \ll 2^{-(r-1/q)l} l^{(d-1)(1-1/\theta)}.$$

Applying Lemma 2.1 with $a = r - 1/q$ and $b = 1 - 1/\theta$ in the case $1 < q \leq 2 \leq p < \infty$, we obtain required upper bounds. The upper bounds in the case $2 \leq q \leq p < \infty$ follow from the above case with $q = 2$. The lower bounds in the case $1 < q \leq 2 \leq p < \infty$ follow from the case $1 < q \leq 2$, $p = 2$.

It was proved in [9] that

$$\sigma_m(\mathbf{H}_\infty^r \cap \mathcal{T}(\Delta Q_n))_p \gg m^{-r} (\log m)^{(d-1)(r+1/2)} \quad (2.18)$$

with some n such that $m \asymp 2^n n^{d-1}$. It is easy to see that for any $f \in \mathbf{H}_\infty^r \cap \mathcal{T}(\Delta Q_n)_p$ we have

$$\|f\|_{\mathbf{B}_{q,\theta}^r} \ll n^{(d-1)/\theta}. \quad (2.19)$$

Relations (2.18) and (2.19) imply the lower bound in the case $2 \leq q \leq p < \infty$. \square

In the case $p = \infty$ we have.

Theorem 2.16. *We have*

$$\sigma_m(\mathbf{H}_{q,\theta}^r)_\infty \ll \begin{cases} m^{-r+\eta} (\log m)^{(d-1)(r-1/q+1-1/\theta)+1/2}, & 1 < q \leq 2, \quad r > 1/q, \\ m^{-r} (\log m)^{(r+1/2-1/\theta)(d-1)+1/2}, & 2 \leq q < \infty, \quad r > 1/2. \end{cases}$$

The upper bounds in Theorem 2.16 are provided by a constructive method $A_m(\cdot, \infty, \mu)$ based on greedy algorithms.

We formulate some known results in the case $1 < p \leq q \leq \infty$.

Theorem 2.17. *Let $1 < p \leq q < \infty$, $r > 0$. Then*

$$\sigma_m(\mathbf{W}_q^r)_p \asymp m^{-r}(\log m)^{(d-1)r}.$$

The upper bound in Theorem 2.17 follows from error bounds for approximation by the hyperbolic cross polynomials (see [16], Ch.2, §2)

$$E_{Q_n}(\mathbf{W}_q^r, L_q) \ll 2^{-rn}, \quad 1 < q < \infty.$$

The lower bound in Theorem 2.17 was proved in [9].

The following result for \mathbf{H}_q^r classes is known.

Theorem 2.18. *Let $p \leq q$, $2 \leq q \leq \infty$, $1 < p < \infty$, $r > 0$. Then*

$$\sigma_m(\mathbf{H}_q^r)_p \asymp m^{-r}(\log m)^{(d-1)(r+1/2)}.$$

The lower bound for all $p > 1$

$$\sigma_m(\mathbf{H}_\infty^r)_p \gg m^{-r}(\log m)^{(d-1)(r+1/2)}$$

was obtained in [9]. The matching upper bounds follow from approximation by the hyperbolic cross polynomials (see [16], Ch.2, Theorem 2.2)

$$E_{Q_n}(\mathbf{H}_q^r)_q := \sup_{f \in \mathbf{H}_q^r} E_{Q_n}(f)_q \asymp n^{(d-1)/2} 2^{-rn}, \quad 2 \leq q < \infty.$$

The following result for \mathbf{B} classes was proved in [12].

Theorem 2.19. *Let $1 < p \leq q < \infty$, $2 \leq q < \infty$, $1 < p < \infty$, $r > 0$. Then*

$$\sigma_m(\mathbf{B}_{q,\theta}^r)_p \asymp m^{-r}(\log m)^{(d-1)(r+1/2-1/\theta)_+}.$$

3 Application of the entropy numbers

Let X be a Banach space and let B_X denote the unit ball of X with the center at 0. Denote by $B_X(y, r)$ a ball with center y and radius r : $\{x \in X : \|x - y\| \leq r\}$. For a compact set A and a positive number ϵ we define the covering number $N_\epsilon(A)$ as follows

$$N_\epsilon(A) := N_\epsilon(A, X) := \min\{n : \exists y^1, \dots, y^n : A \subseteq \cup_{j=1}^n B_X(y^j, \epsilon)\}.$$

For a compact A we define an ϵ -distinguishable set $\{x^1, \dots, x^m\} \subseteq A$ as a set with the property

$$\|x^i - x^j\| > \epsilon, \quad \text{for all } i, j : i \neq j. \quad (3.1)$$

Denote by $M_\epsilon(A) := M_\epsilon(A, X)$ the maximal cardinality of ϵ -distinguishable sets of a compact A . The following simple theorem is well known.

Theorem 3.1. *For any compact set A we have*

$$M_{2\epsilon}(A) \leq N_\epsilon(A) \leq M_\epsilon(A). \quad (3.2)$$

Consider the entropy numbers $\epsilon_k(A, X)$:

$$\epsilon_k(A, X) := \inf\{\epsilon : \exists y^1, \dots, y^{2^k} \in X : A \subseteq \cup_{j=1}^{2^k} B_X(y^j, \epsilon)\}.$$

The following theorem is from [29].

Theorem 3.2. *Let a compact $F \subset X$ be such that there exists a normalized system \mathcal{D} , $|\mathcal{D}| = N$, and a number $r > 0$ such that*

$$\sigma_k(F, \mathcal{D})_X \leq k^{-r}, \quad m \leq N.$$

Then for $k \leq N$

$$\epsilon_k(F, X) \leq C(r) \left(\frac{\log(2N/k)}{k} \right)^r. \quad (3.3)$$

We use the above theorem to prove the following lower bound for best m -term approximations.

Theorem 3.3. *In the case $d = 2$ the following lower bound holds for any $q < \infty$, $r > 1/q$*

$$\sigma_m(\mathbf{W}_q^r)_\infty \gg m^{-r}(\log m)^{1/2}.$$

Proof. We will use a special inequality from [20], which is called the Small Ball Inequality. For an even number n define

$$Y_n := \{\mathbf{s} = (2n_1, 2n_2), \quad n_1 + n_2 = n/2\}.$$

Then for any coefficients $\{c_{\mathbf{k}}\}$

$$\left\| \sum_{\mathbf{s} \in Y_n} \sum_{\mathbf{k} \in \rho(\mathbf{s})} c_{\mathbf{k}} e^{i(\mathbf{k}, \mathbf{x})} \right\|_\infty \geq C \sum_{\mathbf{s} \in Y_n} \left\| \sum_{\mathbf{k} \in \rho(\mathbf{s})} c_{\mathbf{k}} e^{i(\mathbf{k}, \mathbf{x})} \right\|_1, \quad (3.4)$$

where C is a positive number. Inequality (3.4) plays a key role in the proof of lower bounds for the entropy numbers.

Take any even $n \in \mathbb{N}$, which will be chosen, depending on m , later. Consider the following compact

$$F(Y_n)_\infty := \{t = \sum_{\mathbf{s} \in Y_n} t_{\mathbf{s}} : t_{\mathbf{s}} \in \mathcal{T}(\rho(\mathbf{s})), \quad \|t_{\mathbf{s}}\|_\infty \leq 1\}.$$

The known results on volumes of sets of Fourier coefficients of trigonometric polynomials imply the following lemma (see [17] and [20]).

Lemma 3.1. *There exist $2^{n2^{n-1}}$ functions $f_j \in F(Y_n)_\infty$, $j = 1, \dots, 2^{n2^{n-1}}$ such that for $i \neq j$*

$$\|f_i - f_j\|_2 \gg n^{1/2}.$$

We now show that for f_j from Lemma 3.1 we have

$$\|f_i - f_j\|_\infty \gg n.$$

Indeed, for any $f \in F(Y_n)_\infty$ we have

$$\|f\|_2^2 = \sum_{\mathbf{s} \in Y_n} \|t_{\mathbf{s}}\|_2^2 \leq \sum_{\mathbf{s} \in Y_n} \|t_{\mathbf{s}}\|_1 \|t_{\mathbf{s}}\|_\infty \leq \sum_{\mathbf{s} \in Y_n} \|t_{\mathbf{s}}\|_1.$$

It remains to apply the Small Ball Inequality (3.4). Therefore, for $k = n2^{n-1}$ we get by Theorem 3.1

$$\epsilon_k(F(Y_n)_\infty)_\infty \gg \log k.$$

We now use Theorem 3.2. We specify $F := F(Y_n)_\infty$, $\mathcal{D} := \{e^{i(\mathbf{k}, \mathbf{x})} : \|\mathbf{k}\|_\infty \leq 2^{n+1}\}$, $X := L_\infty$. It is clear that for $l \geq \dim T(Y_n) \asymp n2^n \asymp k$ we have

$$\sigma_l(F, \mathcal{D}) = 0.$$

Also, for any $f \in F$ we have

$$\|f\|_\infty \leq n/2 \ll \log k.$$

Denote

$$B := \max_l l^r \sigma_l(F, \mathcal{D})_\infty.$$

By Theorem 3.2 we obtain

$$\log k \ll B n^r k^{-r} \quad \text{and} \quad B \gg n^{-r} k^r \log k.$$

This implies that there is $l \asymp k$ such that

$$\sigma_l(F, \mathcal{D})_\infty \gg n^{-r} \log k \asymp (\log k)^{1-r}.$$

Next, it is clear that for any m

$$\sigma_m(F, \mathcal{D})_\infty \ll \sigma_m(F, \mathcal{T})_\infty.$$

Further, by Littlewood-Paley theorem there is $c_1(q) > 0$ such that

$$c_1(q)n^{-1/2}2^{-rn}F \subset \mathbf{W}_q^r, \quad q < \infty.$$

This completes the proof. □

4 Numerical Integration

4.1 Notations. The problem setting

Numerical integration seeks good ways of approximating an integral

$$\int_{\Omega} f(\mathbf{x}) d\mu$$

by an expression of the form

$$\Lambda(f, X_m) := \sum_{j=1}^m \lambda_j f(\xi^j), \tag{4.1}$$

where $X_m := \{\xi^1, \dots, \xi^m\}$, $\xi^j \in \Omega$, $j = 1, \dots, m$. For a function class \mathbf{W} denote

$$\Lambda(\mathbf{W}, X_m) := \sup_{f \in \mathbf{W}} \left| \int_{\Omega} f(\mathbf{x}) d\mu - \Lambda(f, X_m) \right|.$$

We are interested in dependence on m of the best m -knot error of numerical integration

$$\delta_m(\mathbf{W}) := \inf_{\lambda_1, \dots, \lambda_m; \xi^1, \dots, \xi^m} \Lambda(\mathbf{W}, X_m)$$

for some function classes \mathbf{W} .

4.2 Known lower bounds

The reader can find results and historical comments on numerical integration of classes of functions with mixed smoothness in the book [22], Ch.4 and the survey paper [26]. The following theorem was proved in [18].

Theorem 4.1. *The following lower estimate is valid for any cubature formula (Λ, X_m) with m knots ($r > 1/p$)*

$$\Lambda(\mathbf{W}_p^r, X_m) \geq C(r, d, p) m^{-r} (\log m)^{\frac{d-1}{2}}, \quad 1 \leq p < \infty.$$

The proof of this theorem is based on Theorem 1.8 from Introduction. Theorem 1.8 is used to prove the following assertion.

Lemma 4.1. *Let the coordinates of the vector \mathbf{s} be natural numbers and $\|\mathbf{s}\|_1 = n$. Then for any $N \leq 2^{n-1}$ and an arbitrary cubature formula (Λ, X_N) with N knots there is a $t_{\mathbf{s}} \in \mathcal{T}(2^{\mathbf{s}}, d)$ such that $\|t_{\mathbf{s}}\|_{\infty} \leq 1$ and*

$$\hat{t}_{\mathbf{s}}(0) - \Lambda(t_{\mathbf{s}}, X_N) \geq C(d) > 0. \quad (4.2)$$

For a given m choose n such that

$$m \leq 2^{n-1} < 2m.$$

Consider the polynomial

$$t(\mathbf{x}) = \sum_{\|\mathbf{s}\|_1 = n} t_{\mathbf{s}}(\mathbf{x}),$$

where $t_{\mathbf{s}}$ are polynomials from Lemma 4.1 with $N = m$. Then

$$\hat{t}(0) - \Lambda(t, X_m) \geq C(d) n^{d-1}. \quad (4.3)$$

The proof of Theorem 4.1 was completed by establishing that

$$\|t\|_{\mathbf{W}_p^r} \ll \|t\|_{\mathbf{B}_{p,2}^r} \ll 2^{rn} n^{(d-1)/2}. \quad (4.4)$$

Theorem 4.1 gives the same lower bound for different parameters $1 \leq p < \infty$. It is clear that the bigger the p the stronger the statement.

4.3 New lower bounds

We obtain lower bounds for numerical integration with respect to a special class of knots. Let $\mathbf{s} = (s_1, \dots, s_d)$, $s_j \in \mathbb{N}_0$, $j = 1, \dots, d$. We associate with \mathbf{s} a web $W(\mathbf{s})$ as follows: denote

$$w(\mathbf{s}, \mathbf{x}) := \prod_{j=1}^d \sin(2^{s_j} x_j)$$

and define

$$W(\mathbf{s}) := \{\mathbf{x} : w(\mathbf{s}, \mathbf{x}) = 0\}.$$

Definition 4.1. We say that a set of knots $X_m := \{\xi^i\}_{i=1}^m$ is an (n, l) -net if $|X_m \setminus W(\mathbf{s})| \leq 2^l$ for all \mathbf{s} such that $\|\mathbf{s}\|_1 = n$.

Theorem 4.2. For any cubature formula (Λ, X_m) with respect to a $(n, n-1)$ -net X_m we have

$$\Lambda(\mathbf{W}_p^r, X_m) \gg 2^{-rn} n^{(d-1)/2}, \quad 1 \leq p < \infty.$$

Proof. This proof is similar to that of Theorem 4.1 from [18]. Take \mathbf{s} such that $\|\mathbf{s}\|_1 = n$ and consider $\mathcal{T}(\mathbf{N}, d)$ with $N_j := 2^{s_j-1}$, $j = 1, \dots, d$. Then

$$\dim \mathcal{T}(\mathbf{N}, d) \geq 2^{\|\mathbf{s}\|_1} = 2^n.$$

Let $I(\mathbf{s})$ be a set of indexes such that

$$X_m \setminus W(\mathbf{s}) = \{\xi^i\}_{i \in I(\mathbf{s})}.$$

Then by our assumption $|I(\mathbf{s})| \leq 2^{n-1}$. Consider

$$\Psi(\mathbf{s}) := \{t \in \mathcal{T}(\mathbf{N}, d) : t(\xi_i) = 0, i \in I(\mathbf{s})\}.$$

Then $\dim \Psi(\mathbf{s}) \geq 2^{n-1}$. By Theorem 1.8 we find $t_{\mathbf{s}}^1 \in \Psi(\mathbf{s})$ such that $\|t_{\mathbf{s}}^1\|_{\infty} = 1$ and $\|t_{\mathbf{s}}^1\|_2 \geq c(d) > 0$. Consider

$$t(\mathbf{x}) := \sum_{\|\mathbf{s}\|_1 = n} t_{\mathbf{s}}(\mathbf{x}), \quad t_{\mathbf{s}}(\mathbf{x}) := |t_{\mathbf{s}}^1(\mathbf{x})|^2 w(\mathbf{s}, \mathbf{x})^2.$$

We have

$$t_{\mathbf{s}}^1(\mathbf{x}) w(s_1, x_1)$$

$$= (2i)^{-1} \left(\sum_{\mathbf{k}: |k_j| \leq 2^{s_j-1}} \hat{t}_s^1(\mathbf{k}) e^{i(\mathbf{k}, \mathbf{x}) + i2^{s_1} x_1} + \sum_{\mathbf{k}: |k_j| \leq 2^{s_j-1}} \hat{t}_s^1(\mathbf{k}) e^{i(\mathbf{k}, \mathbf{x}) - i2^{s_1} x_1} \right).$$

and

$$\|t_s^1(\mathbf{x})w(s_1, x_1)\|_2^2 = 2^{-1} \|t_s^1\|_2^2.$$

Therefore,

$$\|t_s^1(\mathbf{x})w(\mathbf{s}, \mathbf{x})\|_2^2 = 2^{-d} \|t_s^1\|_2^2 \geq c_1(d) > 0.$$

Then relation (4.3) is obviously satisfied for our t . Relation (4.4) is proved in the same way as it was proved in [18]. \square

The example that was constructed in the proof of Theorem 4.1 (see above) provides the lower bound for the Besov-type classes. Other proof of Theorem 4.3 is given in [5].

Theorem 4.3. *The following lower estimate is valid for any cubature formula (Λ, X_m) with m knots ($r > 1/p$)*

$$\Lambda(\mathbf{B}_{p,\theta}^r, X_m) \geq C(r, d, p) m^{-r} (\log m)^{(d-1)(1-1/\theta)}, \quad 1 \leq p \leq \infty, \quad 1 \leq \theta \leq \infty.$$

Indeed, the proof of (4.4) from [18] implies

$$\|t\|_{\mathbf{B}_{p,\theta}^r} \ll 2^{rn} n^{(d-1)/\theta}. \quad (4.5)$$

In the same way the proof of Theorem 4.2 gives the following result.

Theorem 4.4. *For any cubature formula (Λ, X_m) with respect to a $(n, n-1)$ -net X_m we have*

$$\Lambda(\mathbf{B}_{p,\theta}^r, X_m) \gg 2^{-rn} n^{(d-1)(1-1/\theta)}, \quad 1 \leq p \leq \infty.$$

We note that Theorems 4.2 and 4.4 provide lower bounds for numerical integration with respect to sparse grids and their modifications. For $n \in \mathbb{N}$ we define the sparse grid $SG(n)$ as follows

$$SG(n) := \{\xi(\mathbf{n}, \mathbf{k}) = (\pi k_1 2^{-n_1}, \dots, \pi k_d 2^{-n_d}),$$

$$0 \leq k_j < 2^{n_j}, j = 1, \dots, d, \quad \|\mathbf{n}\|_1 = n\}.$$

Then it is easy to check that $SG(n) \subset W(\mathbf{s})$ with any \mathbf{s} such that $\|\mathbf{s}\|_1 = n$. Indeed, let $\xi(\mathbf{n}, \mathbf{k}) \in SG(n)$. Take any \mathbf{s} with $\|\mathbf{s}\|_1 = n$. Then $\|\mathbf{s}\|_1 = \|\mathbf{n}\|_1$ and there exists j such that $s_j \geq n_j$. For this j we have

$$\sin 2^{s_j} \xi(\mathbf{n}, \mathbf{k})_j = \sin 2^{s_j} \pi k_j 2^{-n_j} = 0 \quad \text{and} \quad w(\mathbf{s}, \xi(\mathbf{n}, \mathbf{k})) = 0.$$

This means that $SG(n)$ is an (n, l) -net for any l . We note that $|SG(n)| \asymp 2^n n^{d-1}$. It is known (see [15]) that there exists a cubature formula $(\Lambda, SG(n))$ such that

$$\Lambda(\mathbf{H}_p^r, SG(n)) \ll 2^{-rn} n^{d-1}, \quad 1 \leq p \leq \infty, \quad r > 1/p. \quad (4.6)$$

Theorem 4.4 with $\theta = \infty$ shows that the bound (4.6) is sharp. Moreover, Theorem 4.4 shows that even an addition of extra 2^{n-1} arbitrary knots to $SG(n)$ will not improve the bound in (4.6). In the case $X_m = SG(n)$ other proof of Theorem 4.4 is given in [5].

5 Approximate recovery

Consider the following recovering operator. For fixed m , $X_m := \{\xi^j\}_{j=1}^m$, and $\psi_1(\mathbf{x}), \dots, \psi_m(\mathbf{x})$ define the linear operator

$$\Psi(f, X_m) := \sum_{j=1}^m f(\xi^j) \psi_j(\mathbf{x}).$$

For a function class \mathbf{W} define

$$\Psi(\mathbf{W}, X_m)_p := \sup_{f \in \mathbf{W}} \|f - \Psi(f, X_m)\|_p.$$

The main result of this section is the following theorem.

Theorem 5.1. *For any recovering operator $\Psi(\cdot, X_m)$ with respect to a $(n, n-1)$ -net X_m we have for $1 \leq q < p < \infty$*

$$\Psi(\mathbf{H}_q^r, X_m)_p \gg 2^{-n(r-\beta)} n^{(d-1)/p}, \quad \beta := 1/q - 1/p.$$

Before proceeding to the proof of this theorem we make some historical comments. The problem of optimal recovery on classes of functions with mixed smoothness is wide open. Denote for a class \mathbf{W}

$$\varrho_m(\mathbf{W})_p := \inf_{X_m; \psi_1, \dots, \psi_m} \Psi(\mathbf{W}, X_m)_p.$$

The right order of this characteristic is known only in a few cases. It was established in [19] that

$$\varrho_m(\mathbf{W}_2^r)_\infty \asymp m^{-r+1/2}(\log m)^{r(d-1)}. \quad (5.1)$$

The upper bound in (5.1) was obtained by recovering with the Smolyak-type operator T_n with appropriate n . The operator T_n uses the sparse grid $SG(n+d)$ and $\psi_j \in \mathcal{T}(Q_{n+d})$. The operator T_n and its variants were studied in many papers. It is proved in [15] that for any $f \in \mathbf{H}_p^r$, $1 \leq p \leq \infty$, $r > 1/p$

$$\|f - T_n(f)\|_p \ll 2^{-rn} n^{d-1}. \quad (5.2)$$

The following bound is obtained in [22], Ch.4, Section 5, Remark 2: for any $f \in \mathbf{H}_q^r$, $1 \leq q < p < \infty$, $r > 1/q$

$$\|f - T_n(f)\|_p \ll 2^{-n(r-\beta)} n^{(d-1)/p}. \quad (5.3)$$

The upper bound (5.2) and the lower bound for the Kolmogorov width from [21]: for $d = 2$

$$d_m(\mathbf{H}_\infty^r, L_\infty) \asymp m^{-r}(\log m)^{r+1}$$

imply for $d = 2$

$$\varrho_m(\mathbf{H}_\infty^r)_\infty \asymp m^{-r}(\log m)^{r+1}. \quad (5.4)$$

The upper bound (5.3) and known bounds for the Kolmogorov width (see, for instance, [16], Ch.3): for $1 \leq q \leq 2$, $r > 1/q$

$$d_m(\mathbf{H}_q^r, L_2) \asymp m^{-r+\eta}(\log m)^{(d-1)(r+1-1/q)}, \quad \eta := 1/q - 1/2.$$

imply for $1 \leq q \leq 2$, $r > 1/q$

$$\varrho_m(\mathbf{H}_q^r)_2 \asymp m^{-r+\eta}(\log m)^{(d-1)(r+1-1/q)}. \quad (5.5)$$

As we already pointed out above $T_n(f) \in \mathcal{T}(Q_{n+d})$. Thus T_n provides approximation from the hyperbolic cross $\mathcal{T}(Q_{n+d})$. The upper bound (5.3) and known bounds for the best hyperbolic cross approximation (see, for instance, [16], Ch.2, Theorem 2.2) show that for $1 \leq q < p < \infty$, $r > 1/q$, the operator T_n provides optimal in the sense of order rate of approximation for \mathbf{H}_q^r in the L_p .

Proof of Theorem 5.1. We use the polynomials $t_{\mathbf{s}}^1$ constructed in the proof of Theorem 4.2. We also need some more constructions. Let

$$\mathcal{K}_N(x) := \sum_{|k| \leq N} (1 - |k|/N) e^{ikx} = (\sin(Nx/2))^2 / (N(\sin(x/2))^2)$$

be a univariate Fejér kernel. The Fejér kernel \mathcal{K}_N is an even nonnegative trigonometric polynomial in $\mathcal{T}(N-1)$. From the obvious relations

$$\|\mathcal{K}_N\|_1 = 1, \quad \|\mathcal{K}_N\|_\infty = N$$

and the inequality

$$\|f\|_q \leq \|f\|_1^{1/q} \|f\|_\infty^{1-1/q}$$

we get

$$CN^{1-1/q} \leq \|\mathcal{K}_N\|_q \leq N^{1-1/q}, \quad 1 \leq q \leq \infty. \quad (5.6)$$

In the multivariate case define

$$\mathcal{K}_{\mathbf{N}}(\mathbf{x}) := \prod_{j=1}^d \mathcal{K}_{N_j}(x_j), \quad \mathbf{N} = (N_1, \dots, N_d).$$

Then the $\mathcal{K}_{\mathbf{N}}$ are nonnegative trigonometric polynomials from $\mathcal{T}(\mathbf{N} - \mathbf{1}, d)$ which have the following properties:

$$\|\mathcal{K}_{\mathbf{N}}\|_1 = 1, \quad (5.7)$$

$$\|\mathcal{K}_{\mathbf{N}}\|_q \asymp \vartheta(\mathbf{N})^{1-1/q}, \quad 1 \leq q \leq \infty. \quad (5.8)$$

For n of the form $n = 4l$, $l \in \mathbb{N}$, define

$$Y(n, d) := \{\mathbf{s} : \mathbf{s} = (4l_1, \dots, 4l_d), \quad l_1 + \dots + l_d = n/4, \quad l_j \in \mathbb{N}, \quad j = 1, \dots, d\}.$$

Define for $\mathbf{s} \in Y(n, d)$

$$t_{\mathbf{s}}(\mathbf{x}) := t_{\mathbf{s}}^1(\mathbf{x}) \mathcal{K}_{2^{\mathbf{s}-2}}(\mathbf{x} - \mathbf{x}^*),$$

where \mathbf{x}^* is a point of maximum of $|t_{\mathbf{s}}^1(\mathbf{x})|$. Finally, define

$$t(\mathbf{x}) := \sum_{\mathbf{s} \in Y(n, d)} t_{\mathbf{s}}(\mathbf{x}) w(\mathbf{s}, \mathbf{x}).$$

Then we have

$$|t_{\mathbf{s}}(\mathbf{x}^*)| \gg 2^n$$

and, therefore, by Nikol'skii's inequality

$$\|t_{\mathbf{s}}\|_2 \gg 2^{n/2}.$$

It follows from our definition of $Y(n, d)$ that polynomials $t_{\mathbf{s}}(\mathbf{x})w(\mathbf{s}, \mathbf{x})$, $\mathbf{s} \in Y(n, d)$, form an orthogonal system. This implies

$$\|t\|_2^2 \gg 2^n n^{d-1}. \quad (5.9)$$

Take any $p \in (1, \infty)$ and by Theorem 1.6 estimate

$$\|t\|_{p'} \ll \left(\sum_{\mathbf{s} \in Y(n, d)} \|t_{\mathbf{s}}\|_1^{p'} 2^{\|\mathbf{s}\|_1(p'-1)} \right)^{1/p'} \ll 2^{n/p} n^{(d-1)/p'}. \quad (5.10)$$

Relations (5.9) and (5.10) imply

$$\|t\|_p \gg 2^{n(1-1/p)} n^{(d-1)/p}. \quad (5.11)$$

It is clear that

$$\|t\|_{\mathbf{H}_q^r} \ll 2^{n(r+1-1/q)}. \quad (5.12)$$

Bounds (5.11) and (5.12) imply the required in Theorem 5.1 bound. Theorem 5.1 is proved.

The inequality

$$\|t\|_{\mathbf{B}_{q,\theta}^r} \ll 2^{n(r+1-1/q)} n^{(d-1)/\theta}$$

and (5.11) imply the following result.

Theorem 5.2. *For any recovering operator $\Psi(\cdot, X_m)$ with respect to a $(n, n-1)$ -net X_m we have for $1 \leq q < p < \infty$, $r > \beta$,*

$$\Psi(\mathbf{B}_{q,\theta}^r, X_m)_p \gg 2^{-n(r-\beta)} n^{(d-1)(1/p-1/\theta)}, \quad \beta := 1/q - 1/p.$$

6 Discussion

As we stressed in the title and in the Introduction we are interested in constructive methods of m -term approximation with respect to the trigonometric system. Theorem 1.1, basically, solves this problem for approximation in L_p , $2 \leq p < \infty$. We do not have a similar result for approximation in L_p , $1 < p < 2$. The corresponding Lebesgue-type inequality from [30] gives for $1 < p < 2$

$$\|f_{C(t,p,d)m^{p'-1}\log(m+1)}\|_p \leq C\sigma_m(f)_p,$$

which is much weaker than Theorem 1.1. It would be interesting to obtain good Lebesgue-type inequalities in the case $1 < p < 2$ for either the WCGA or for some other constructive methods.

Main results of this paper are on the m -term approximation in the case $2 \leq p \leq \infty$. For $p \in [2, \infty)$ the situation is very good: we have a universal algorithm (WCGA), which provides almost optimal (up to extra $(\log m)^{C(r,d)}$ factor) m -term approximation for all classes \mathbf{W}_q^r , \mathbf{H}_q^r , and $\mathbf{B}_{q,\theta}^r$. Also, there are constructive methods, based on greedy algorithms, which provide the optimal rate for the above classes. However, the upper bounds in, say, Theorem 1.4 hold for smoothness r larger than the one required for embedding of \mathbf{W}_q^r into L_p . It would be interesting to find constructive methods, which provide right orders of decay of $\sigma_m(\mathbf{W}_q^r)_p$, $\sigma_m(\mathbf{H}_q^r)_p$, and $\sigma_m(\mathbf{B}_{q,\theta}^r)_p$ for small smoothness.

The case $p = \infty$ (approximation in the uniform norm) is a very interesting and difficult case. The space $C(\mathbb{T}^d)$ is not a smooth Banach space. Therefore, the existing greedy approximation theory does not apply directly in the case of approximation in L_∞ . In particular, there is no analog of Theorem 1.1 in the case $p = \infty$. However, for the function classes with mixed smoothness there is a way around this problem. As it is demonstrated in the proof of Theorem 2.4 we can use greedy algorithms in L_p with large p to obtain bounds on m -term approximation in L_∞ . The price we pay for this trick is an extra $(\log m)^{1/2}$ factor in the error bound. This extra factor results from the factor $p^{1/2}$ in the error bounds of approximation by greedy algorithms in L_p , $2 \leq p < \infty$ (see Remark 2.1 and Theorem 2.2). An extra $(\log m)^{1/2}$ appears, as a result of different techniques, in other upper bounds of asymptotic characteristics of classes of functions with mixed smoothness, when we go from $p < \infty$ to $p = \infty$ (see, for instance, [28], Ch.3, Section 3.6). Unfortunately, we do not have matching lower bounds for our upper bounds

for m -term approximation in L_∞ . A very special case in Theorem 3.3 could be interpreted as a hint that we cannot get rid of that extra $(\log m)^{1/2}$ for approximation in L_∞ .

We discussed isotropic classes of functions with mixed smoothness. Isotropic means that all variables play the same role in the definition of our smoothness classes. In the hyperbolic cross approximation theory anisotropic classes of functions with mixed smoothness are of interest and importance. We give the corresponding definitions. Let $\mathbf{r} = (r_1, \dots, r_d)$ be such that $0 < r_1 = r_2 = \dots = r_\nu < r_{\nu+1} \leq r_{\nu+2} \leq \dots \leq r_d$ with $1 \leq \nu \leq d$. For $\mathbf{x} = (x_1, \dots, x_d)$ denote

$$F_{\mathbf{r}}(\mathbf{x}) := \prod_{j=1}^d F_{r_j}(x_j)$$

and

$$\mathbf{W}_p^{\mathbf{r}} := \{f : f = \varphi * F_{\mathbf{r}}, \quad \|\varphi\|_p \leq 1\}.$$

We now proceed to classes $\mathbf{H}_q^{\mathbf{r}}$ and $\mathbf{B}_{q,\theta}^{\mathbf{r}}$. Define

$$\|f\|_{\mathbf{H}_q^{\mathbf{r}}} := \sup_{\mathbf{s}} \|\delta_{\mathbf{s}}(f)\|_q 2^{(\mathbf{r}, \mathbf{s})},$$

and for $1 \leq \theta < \infty$ define

$$\|f\|_{\mathbf{B}_{q,\theta}^{\mathbf{r}}} := \left(\sum_{\mathbf{s}} (\|\delta_{\mathbf{s}}(f)\|_q 2^{(\mathbf{r}, \mathbf{s})})^\theta \right)^{1/\theta}.$$

We will write $\mathbf{B}_{q,\infty}^{\mathbf{r}} := \mathbf{H}_q^{\mathbf{r}}$. Denote the corresponding unit ball

$$\mathbf{B}_{q,\theta}^{\mathbf{r}} := \{f : \|f\|_{\mathbf{B}_{q,\theta}^{\mathbf{r}}} \leq 1\}.$$

It is known that in many problems of estimating asymptotic characteristics the anisotropic classes of functions of d variables with mixed smoothness behave in the same way as isotropic classes of functions of ν variables (see, for instance, [16]). It is clear that the above remark holds for the lower bounds. To prove it for the upper bounds one needs to develop, in some cases, a special technique. The techniques developed in this paper work for the anisotropic classes as well. For instance, the main Lemma 2.1 is replaced by the following lemma.

Lemma 6.1. Denote $r := r_1$. Define for $f \in L_1$

$$f_{l,r} := \sum_{\mathbf{s}: r_l \leq (\mathbf{r}, \mathbf{s}) < r(l+1)} \delta_{\mathbf{s}}(f), \quad l \in \mathbb{N}_0.$$

Consider the class

$$\mathbf{W}_A^{\mathbf{r}, a, b} := \{f : \|f_{l,r}\|_A \leq 2^{-al} l^{(\nu-1)b}\}.$$

Then for $2 \leq p < \infty$ and $a > 0$ there is a constructive method based on greedy algorithms, which provides the bound

$$\sigma_m(\mathbf{W}_A^{\mathbf{r}, a, b})_p \ll m^{-a-1/2} (\log m)^{(\nu-1)(a+b)}. \quad (6.1)$$

For $p = \infty$ we have

$$\sigma_m(\mathbf{W}_A^{\mathbf{r}, a, b})_\infty \ll m^{-a-1/2} (\log m)^{(\nu-1)(a+b)+1/2}. \quad (6.2)$$

Proposition 6.1. Results of Section 2 hold for the anisotropic classes of functions with mixed smoothness with $r = r_1$ and d replaced by ν .

References

- [1] E. S. Belinskii, Approximation by a "floating" system of exponentials on classes of smooth periodic functions, *Matem. Sb.* **132** (1987), 20–27; English translation in *Math. USSR Sb.* **60** (1988).
- [2] E.S. Belinskii, Decomposition theorems and approximation by a "floating" system of exponentials, *Transactions of the American Mathematical Society*, **350** (1998), 43–53.
- [3] R.A. DeVore and V.N. Temlyakov, Nonlinear approximation by trigonometric sums, *J. Fourier Analysis and Applications*, **2** (1995), 29–48.
- [4] S.J. Dilworth, D. Kutzarova, V.N. Temlyakov, Convergence of some Greedy Algorithms in Banach spaces, *The J. Fourier Analysis and Applications* **8** (2002), 489–505.
- [5] Dinh Dung and T. Ullrich, Lower bounds for the integration error for multivariate functions with mixed smoothness and optimal Fibonacci cubature for functions on the square, manuscript, 2014.

- [6] M. Donahue, L. Gurvits, C. Darken and E. Sontag, Rate of convex approximation in non-Hilbert spaces, *Constructive Approx.*, **13** (1997), 187–220.
- [7] E.D. Gluskin, Extremal properties of orthogonal parallelepipeds and their application to the geometry of Banach spaces, *Math USSR Sbornik* **64** (1989), 85–96.
- [8] R.S. Ismagilov, Widths of sets in normed linear spaces and the approximation of functions by trigonometric polynomials, *Uspekhi Mat. Nauk*, **29** (1974), 161–178; English transl. in *Russian Math. Surveys*, **29** (1974).
- [9] B.S. Kashin and V.N. Temlyakov, On best m -term approximations and the entropy of sets in the space L^1 , *Math. Notes*, **56** (1994), 1137–1157.
- [10] V.E. Maiorov, Trigonometric diameters of the Sobolev classes W_p^r in the space L_q , *Math. Notes* **40** (1986), 590–597.
- [11] Y. Makovoz, On trigonometric n -widths and their generalizations, *J. Approx. Theory* **41** (1984), 361–366.
- [12] A.S. Romanyuk, Best M -term trigonometric approximations of Besov classes of periodic functions of several variables, *Izvestia RAN, Ser. Mat.* **67** (2003), 61–100; English transl. in *Izvestiya: Mathematics* (2003), 67(2):265.
- [13] S. B. Stechkin, On the best approximation of given classes of functions by arbitrary polynomials, *Uspekhi Mat. Nauk* **9** (1954), 133–134 (in Russian).
- [14] V.N. Temlyakov, Approximation of Periodic Functions of Several Variables by Bilinear Forms, *Izvestiya AN SSSR*, **50** (1986), 137–155; English transl. in *Math. USSR Izvestija*, **28** (1987), 133–150.
- [15] V.N. Temlyakov, Approximate recovery of periodic functions of several variables, *Mat. Sb.*, **128** (1985), 256–268; English transl. in *Math. USSR-Sb* **56** (1987), 249–261.

- [16] V.N. Temlyakov, Approximation of functions with bounded mixed derivative, *Trudy MIAN*, **178** (1986), 1–112. English transl. in *Proc. Steklov Inst. Math.*, **1** (1989).
- [17] V.N. Temlyakov, Estimates of the asymptotic characteristics of classes of functions with bounded mixed derivative or difference, *Trudy Matem. Inst. Steklov*, **189** (1989), 138–168; English transl. in *Proceedings of the Steklov Institute of Mathematics*, 1990, Issue 4, 161–197.
- [18] V.N. Temlyakov, On a way of obtaining lower estimates for the errors of quadrature formulas, *Matem. Sbornik*, **181** (1990), 1403–1413; English transl. in *Math. USSR Sbornik*, **71** (1992).
- [19] V.N. Temlyakov, On Approximate Recovery of Functions with Bounded Mixed Derivative, *J. Complexity*, **9** (1993), 41–59.
- [20] V.N. Temlyakov, An inequality for trigonometric polynomials and its application for estimating the entropy numbers, *J. Complexity*, **11** (1995), 293–307.
- [21] V.N. Temlyakov, An inequality for trigonometric polynomials and its application for estimating the Kolmogorov widths, *East J. Approximations*, **2** (1996), 253–262.
- [22] V.N. Temlyakov, *Approximation of periodic functions*, Nova Science Publishes, Inc., New York., 1993.
- [23] V.N. Temlyakov, Nonlinear Kolmogorov’s widths, *Matem. Zametki*, **63** (1998), 891–902.
- [24] V.N. Temlyakov, Weak greedy algorithms, *Advances in Comput. Math.* **12** (2000), 213–227.
- [25] V.N. Temlyakov, Greedy algorithms in Banach spaces, *Advances in Comput. Math.* **14** (2001), 277–292.
- [26] V.N. Temlyakov, Cubature formulas and related questions, *J. Complexity* **19** (2003), 352–391.
- [27] V.N. Temlyakov, Greedy-Type Approximation in Banach Spaces and Applications, *Constr. Approx.* **21** (2005), 257–292.

- [28] V.N. Temlyakov, Greedy approximation, Cambridge University Press, 2011.
- [29] V.N. Temlyakov, An inequality for the entropy numbers and its application, J. Approx. Theory, **173** (2013), 110–121.
- [30] V.N. Temlyakov, Sparse approximation and recovery by greedy algorithms in Banach spaces, Forum of Mathematics, Sigma, **2** (2014), e12, 26 pages; IMI Preprint, 2013:09, 1–27; arXiv:1303.6811v1, 27 Mar 2013.